Narratives, Not Hard Restrictions Identifying Oil-Market and Monetary Policy Shocks

Justin Eloriaga* Ana María Herrera† Elena Pesavento‡

July 28, 2026

Abstract

Set-identified structural VARs are routinely disciplined with hard point restrictions whose precise validity cannot be verified. Oil-market models cap the short-run supply elasticity and monetary models impose exact zeros in the policy rule. We show that a small number of narrative restrictions anchored to well-documented historical episodes can replace both devices. In a four-variable oil-market VAR, narrative restrictions reproduce the results of Kilian and Murphy (2014) without the elasticity bound. In a six-variable monetary VAR, they reproduce the benchmark of Arias, Caldara, and Rubio-Ramírez (2019) without the zero restrictions. History, rather than hard restrictions, disciplines the identified set.

JEL Classification. C32, E52, Q43, C11

Keywords. Narrative sign restrictions, SVAR, identification, oil supply elasticity, systematic component of monetary policy

1 Introduction

Since the late 1970s, large variations in oil prices have fueled the interest of policy makers and academics in understanding the dynamic effect of unexpected changes in oil prices on economic outcomes. To estimate these dynamic effects, researchers commonly use vector

**Corresponding Author*. International Monetary Fund, Shanghai Center. Email jeloriaga@imf.org. The views expressed in this paper are those of the authors and do not necessarily represent the views of the IMF, its Executive Board, or IMF Management. We are thankful for the comments of Juan Rubio Ramírez.

†Department of Economics, University of Kentucky. Email amherrera@uky.edu

‡Department of Economics, Emory University. Email epesave@emory.edu

autoregressive (VAR) models of the global oil market. Following Kilian (2009), researchers recognized that disentangling oil supply and demand shocks was crucial to understanding the dynamic response of real oil prices and macroeconomic outcomes and to assessing the historical contribution of various shocks. Since then, several identification strategies such as short-run restrictions, sign restrictions, and combinations of both have been employed to unravel the role of supply and demand shocks.

The evolution of these identification strategies is instructive. The first generation of oil-market VAR models that separately identified demand and supply shocks relied primarily on exclusion restrictions, with the influential model of Kilian (2009) imposing a recursive ordering that placed the oil supply shock first and thereby implied a vertical short-run supply curve. Noting that these short-run exclusion restrictions represented strong a priori assumptions about the prevailing market structure, a second generation of models followed an alternative approach and imposed sign restrictions on the impact multiplier matrix, as in Baumeister and Peersman (2013). However, Kilian and Murphy (2012) demonstrated that agnostic sign restrictions alone were insufficient to identify the response of the real oil price to structural shocks, because purely sign-identified models admit structurally implausible parameter values. To address this issue, Kilian and Murphy (2014) imposed an upper bound of 0.0258 on the impact price elasticity of oil supply when investigating the role of inventories and speculative demand in oil markets.

These elasticity bound restrictions, while providing the additional discipline needed for identification, have been subject to criticism. Baumeister and Hamilton (2019) contended that the bound was too restrictive and questioned whether researchers could know the supply elasticity with such precision, and the meta-analysis of Caldara, Cavallo, and Iacoviello (2019) illustrates that the range of estimates found in the literature is wide.[1] In brief, while it has been shown that elasticity bounds aid in reducing the set of admissible models and lead to more informative impulse responses, no consensus has been reached regarding the appropriate bound, hence the benefit of a method that allows the researcher to circumvent such a choice.

The monetary policy literature contains a parallel debate about a parallel device. Under the agnostic sign restrictions of Uhlig (2005), which constrain the impulse responses of prices, commodity prices, nonborrowed reserves, and the federal funds rate while deliberately leaving output unrestricted, the output response to a contractionary monetary policy shock is ambiguous and mostly positive. Arias, Caldara, and Rubio-Ramírez (2019), henceforth ACRR, overturn this result by restricting the systematic component of monetary policy in-

[1] See Herrera and Rangaraju (2020) for a discussion and analysis of different specifications. Estimates for the supply elasticity in the Caldara, Cavallo, and Iacoviello (2019) meta-analysis range from -0.16 to 0.27.

stead of the impulse responses. Their identification requires the federal funds rate to respond positively to output and prices, and it also imposes two exact zero restrictions requiring that the funds rate not react contemporaneously to total or nonborrowed reserves. The zeros are the monetary analogue of the elasticity bound in the sense that both are hard point restrictions on structural parameters whose precise validity cannot be verified, and both do identifying work that is not visible in the published impulse responses.

This paper demonstrates that narrative sign restrictions, introduced by Antolín-Díaz and Rubio-Ramírez (2018), allow the researcher to circumvent both devices. Narrative restrictions strengthen sign-restriction identification by incorporating historical information about specific episodes, constraining either the sign of structural shocks during well-documented events or the relative contribution of shocks to historical decompositions. Our design is the same in both applications, and it is deliberately simple. We estimate the published benchmark with its hard restriction, we estimate the same model with the narrative added, and we then remove the hard restriction while keeping the narrative. If the last model reproduces the first, the hard restriction is redundant once the narrative record is used.

We make three contributions. First, in the four-variable oil-market VAR of Kilian and Murphy (2014), we show that narrative restrictions anchored to three well-documented supply disruptions render the elasticity bound unnecessary. Impulse responses, variance decompositions, and historical decompositions without the bound are virtually identical to those with it, and the narrative eliminates the implausibly elastic rotations that sign restrictions alone admit. Second, in the six-variable monetary VAR of ACRR, we show that narrative restrictions render the zero restrictions unnecessary. Starting from Uhlig's identification and adding five signed episodes, two qualitative rule signs, and a new restriction type that we call systematic dominance, the coefficients of the policy rule on reserves concentrate at zero without ever being restricted, the rule coefficients on output and prices match ACRR's, and the output puzzle is resolved. The systematic-dominance restriction states that at certain well-documented dates the funds rate was moving with the rule rather than away from it. To our knowledge it is the first narrative restriction imposed on the systematic component of a policy rule rather than on a shock, and we show that the two kinds of narrative information are complements, because restrictions on shocks can only bound the systematic response from above while restrictions on the rule push it up. Third, the two applications together make a general methodological point. The choice between weak identification through sign restrictions and potentially over-restrictive identification through hard point restrictions presents a false dichotomy, and narrative restrictions offer a middle ground. They provide strong identifying discipline without requiring researchers to impose precise numerical bounds or exact exclusions, as long as there exist well-documented historical episodes where economic theory provides clear predictions about shock behavior.

The remainder of the paper proceeds as follows. Section 2 reviews the methodology and introduces the systematic-dominance restriction. Section 3 presents the oil-market application and Section 4 presents the monetary application. Section 5 concludes with practical guidance. Variance decompositions, historical decompositions, sensitivity exercises, and full narrative documentation for every episode appear in the Supplemental Appendix.

2 Methodology

The structural VAR is

$$\mathbf{y}_t'\mathbf{A}_0 = \mathbf{x}_t'\mathbf{A}_+ + \varepsilon_t' \quad \text{for } t = 1, \ldots, T, \tag{1}$$

where $\mathbf{y}_t$ collects the n endogenous variables, $\mathbf{x}_t$ collects p lags and deterministic terms, and the structural innovations ε_t are Gaussian with identity covariance. The reduced-form innovations satisfy $\mathbf{u}_t' = \varepsilon_t'\mathbf{A}_0^{-1}$, so that the reduced-form parameters can be estimated directly while the structural parameters require identification restrictions. Identification through sign restrictions is set identification. We use the term rotation set to mean the collection of orthogonal rotations Q of the reduced-form Cholesky factor that satisfy the maintained restrictions, any rotation in the set is admissible, and posterior summaries are taken across the set. Traditional sign restrictions take the form $\mathbf{S}_j\mathbf{F}(\Theta)\mathbf{e}_{j,n} > \mathbf{0}$, where $\mathbf{F}(\Theta)$ stacks the objects being restricted, which may be impulse responses as in Uhlig (2005) or structural parameters as in ACRR.

Narrative restrictions operate by shrinking the rotation set, discarding rotations whose implied structural shocks or historical decompositions are inconsistent with the narrative evidence at targeted dates. Antolín-Díaz and Rubio-Ramírez (2018) introduce two classes. The first constrains the sign of a structural shock during a documented episode, requiring for instance $\mathbf{e}_{j,n}'\varepsilon_{t_\nu} > 0$ at the anchor date t_ν. The second constrains the historical decomposition, with Type A restrictions requiring the target shock to be the most (or the least) important contributor to the unexpected change in a designated variable and Type B restrictions requiring its absolute contribution to exceed the combined contribution of all other shocks. Because conditioning on events at specific dates distorts the likelihood, draws that survive the narrative restrictions are re-weighted by the inverse of the probability that independent standard normal shocks would satisfy them, exactly as Antolín-Díaz and Rubio-Ramírez (2018) prescribe.

The monetary application adds a third class that restricts the systematic component itself. Write the policy equation of the monetary VAR as a rule, so that the funds-rate innovation decomposes into the rule's contemporaneous response to incoming data and the policy shock,

$$\varepsilon_t^{MP} = a\big[\, u_{\mathrm{FFR},t} - \psi' \tilde{u}_t \,\big], \tag{2}$$

where $\tilde{u}_t$ collects the reduced-form innovations of the non-policy variables, ψ collects the rule coefficients, and a is a scale. A shock-sign restriction at date τ asserts that policy moved more than the rule prescribes at that date. A *systematic dominance* restriction asserts the opposite, namely that the funds-rate movement was mostly the rule responding to the observed data, which we formalize as $|\varepsilon_\tau^{MP}/a| < |\psi' \tilde{u}_\tau|$ at the anchor date. Section 4 shows that the two classes are complements rather than substitutes, because shock-based narratives bound the systematic response from above while systematic-dominance narratives push it up.

3 Application I. The Global Oil Market

We estimate a four-variable model with monthly oil production growth, the Kilian and Zhou (2018) real economic activity index, the real price of oil, and the change in oil inventories over January 1974 to December 2017, following Kilian and Murphy (2014), which refines Kilian and Murphy (2012). Three shocks are named, these being flow supply, flow demand, and speculative (precautionary) demand. We compare three identification schemes throughout the application. The Bound scheme imposes the Kilian and Murphy (2014) sign restrictions together with their elasticity bound and serves as the baseline. The Bound-and-narrative scheme adds the full narrative to the baseline, and the Narrative scheme imposes the sign restrictions and the full narrative but removes the bound. Because the last two schemes differ only in the elasticity bound, their comparison isolates exactly what the bound contributes. Panel A of Table 1 lists the restrictions and Panel A of Table 2 lists the anchor episodes.

Table 1: Identification schemes in the two applications

Panel A. Oil market (Kilian and Murphy, 2014)

Restriction	Bound	Bound and narr.	Narrative
Impact sign restrictions on the three shocks	✓	✓	✓
Dynamic sign restrictions, flow supply ($h = 1, \ldots, 12$)	✓	✓	✓
Elasticity bound, impact supply elasticity $\in (0, 0.0258)$	✓	✓	
Narrative restrictions NSR-1 to NSR-5 at each anchor		✓	✓

Panel B. Monetary policy

Restriction	Uhlig	ACRR	Narrative
IRF signs, prices, comm. prices, NBR fall, FFR rises ($h = 0, \ldots, 5$)	✓		✓
FFR rises on impact	✓	✓	✓
Rule signs $\psi_y \geq 0$ and $\psi_p \geq 0$		✓	✓
Zero restrictions $\psi_{tr} = \psi_{nbr} = 0$		✓	
Signed policy shocks at five dated episodes			✓
Systematic dominance at two dated episodes			✓

Notes. A check mark indicates the restriction is imposed. In Panel A the impact signs follow Kilian and Murphy (2014) and NSR-1 to NSR-5 are described in the text. In Panel B, ψ_x denotes the contemporaneous response of the federal funds rate to variable x in the policy rule. Surviving draws are re-weighted following Antolín-Díaz and Rubio-Ramírez (2018) in both applications.

The narrative consists of five restrictions imposed at each anchor date. The flow-supply shock is an adverse disruption (NSR-1) and is the most important contributor to the oil-production surprise (NSR-2), while the speculative-demand shock is positive (NSR-3), is the most important contributor to inventories (NSR-4), and is the least important contributor to production (NSR-5). The three anchors are the Iranian Revolution of January 1979, the Persian Gulf War of August 1990, and the Venezuelan general strike of December 2002. The defensibility of the restrictions rests on the fact that each anchor pairs an unambiguous physical supply disruption with a documented precautionary scramble for oil, so that the economics of the episode is not in dispute. Strikes in the Iranian oil sector and the revolution that drove the Shah into exile cut Iranian output from roughly five and a half to under one million barrels per day, the invasion of Kuwait embargoed about seven percent of world output, and the strike at the state oil company PDVSA collapsed Venezuelan production from three million to roughly seven hundred thousand barrels per day within weeks. In each case the contemporaneous production decline was the physical shortfall itself, which is what

NSR-1 and NSR-2 assert, and nothing more.

August 1990 illustrates why the precautionary restrictions are equally grounded. The real price of oil nearly doubled between August and October, far more than the physical shortfall warranted once Saudi Arabia and others raised output by some three million barrels per day, and the price then collapsed in January 1991 when Operation Desert Storm resolved the uncertainty, before any improvement in fundamentals. A price spike driven by fear of future scarcity that unwinds when the uncertainty rather than the shortage is resolved is the definition of precautionary demand, and it is precisely how Kilian (2009) reads the episode. The same pattern of fear amplifying a physical loss holds at the other anchors, where the Iranian revolution left the Iran-Iraq War looming and the Venezuelan strike coincided with the build-up to the invasion of Iraq. The restrictions NSR-3 to NSR-5 are also the narrative counterpart of the elasticity bound. In the Kilian and Murphy (2014) architecture the speculative shock is the demand for above-ground inventories, the only channel through which expected future scarcity reaches the current price (Kilian and Lee, 2014), so requiring that shock to own the inventory movement imposes the model's own logic at the dates where it is least controversial. Requiring the same shock to be the least important driver of production (NSR-5) reflects the fact that the production declines were embargoes and strikes rather than responses to speculation and that short-run oil supply is highly price-inelastic. NSR-5 therefore re-imposes the economic content of the bound at the anchor dates without asserting a numerical cap everywhere in the sample.

Figure 1 reports the impulse responses under the multi-anchor design. The Narrative scheme tracks the bounded baseline across all four variables and all three shocks, and the 68 percent credible bands overlap throughout. An adverse flow-supply shock raises the real price persistently while depressing production and activity, a positive flow-demand shock raises activity, the price, and production with a lag, and a positive speculative shock raises the price and inventories while modestly depressing activity. The intermediate scheme behaves as a hybrid, moving with the Narrative scheme in the responses that the narrative pins down while tracking the baseline where the elasticity bound is the operative constraint. The variance and historical decompositions, reported in the Supplemental Appendix, deliver the same conclusion, since the schemes allocate variance almost identically and both attribute the August 1990 production surprise almost entirely to the flow-supply shock.

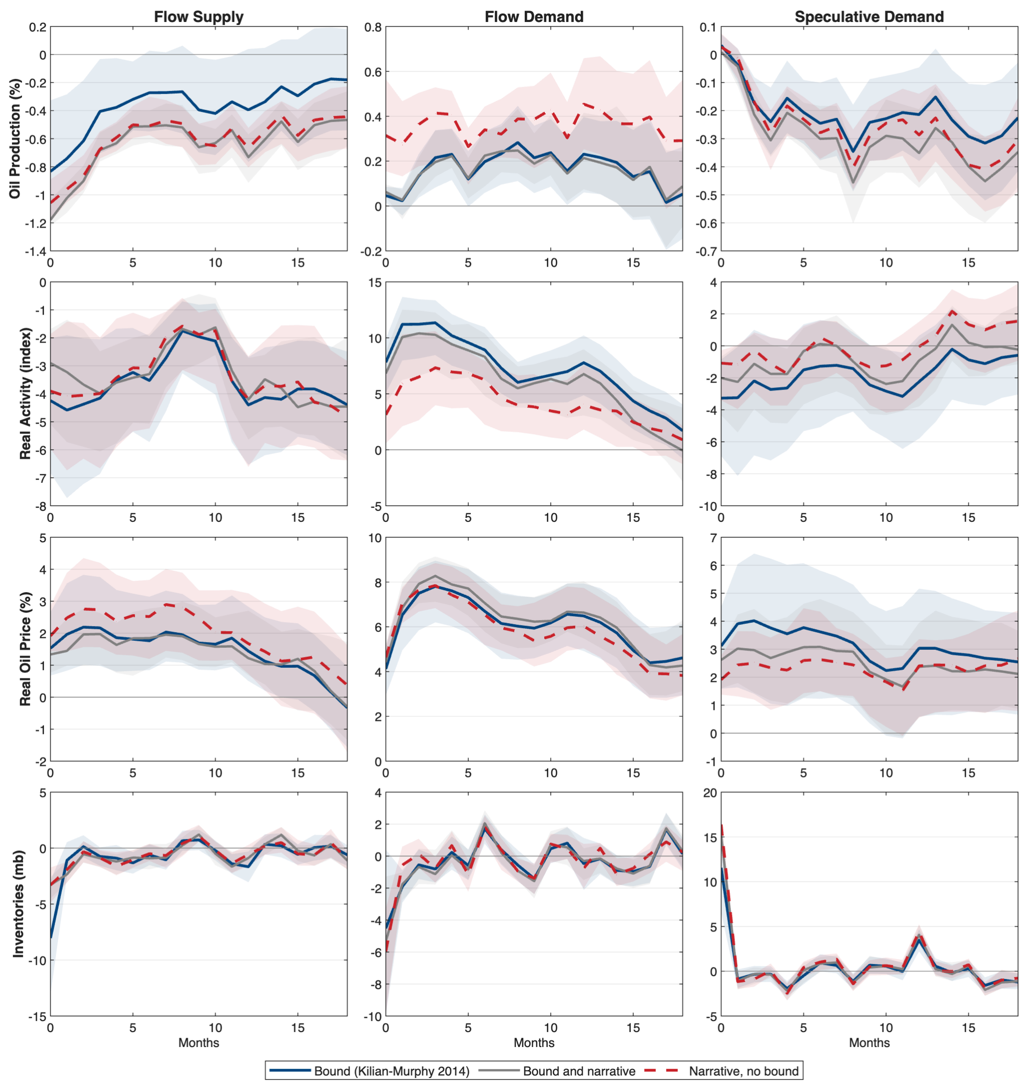

Figure 1: Oil market. Impulse responses under the three identification schemes

Notes. The Bound scheme imposes the sign restrictions and the elasticity bound (blue solid), the Bound-and-narrative scheme adds the narrative (gray solid), and the Narrative scheme imposes the narrative with no bound (red dashed). The narrative is imposed at the three anchors of Table 2. Shaded regions are 68 percent pointwise credible bands.

Two further results complete the picture. First, the narrative does the disciplining work that motivated the bound. Under the sign restrictions alone the implied impact supply elasticity is left almost entirely unpinned, with posterior medians of 0.11 and 0.45 for the two demand shocks and with 85 and 97 percent of admissible rotations exceeding the Kilian and Murphy (2014) bound, which is exactly the configuration Kilian and Murphy (2012) warn against. With the full narrative imposed, the implausibly elastic rotations are eliminated and the implied elasticity settles at small values, tightening toward the baseline as anchors are added. Second, the identification is tight. Imposing the fifteen narrative restrictions across the three events accepts only on the order of tens of draws from a base of six hundred thousand, and this tightness is not a numerical defect but a measure of how sharply the historically grounded restrictions identify the model. Consistent with this reading, when the bound is left in place but relaxed to a non-binding level, the estimates barely move, because by the time the narratives are imposed the rotation set is already disciplined.

4 Application II. The Systematic Component of Monetary Policy

The monetary application uses the six-variable VAR of ACRR with monthly real GDP, the GDP deflator, a commodity price index, total reserves, nonborrowed reserves, and the federal funds rate over 1965 to 2007, with twelve lags and a flat prior. Our replication of their benchmark is exact in the sense that we port their code line by line with the same seed, and our posterior medians for the rule coefficients are 0.848 on output and 2.775 on prices against their published 0.84 and 2.73.

Under Uhlig (2005) the monetary shock is identified by sign restrictions on impulse responses alone, requiring prices, commodity prices, and nonborrowed reserves to fall and the funds rate to rise for six months while leaving the policy rule free. Under ACRR the rule itself is restricted, with the funds rate responding positively to output and prices, written $\psi_y \geq 0$ and $\psi_p \geq 0$, and with the two reserves coefficients set to zero exactly. The two identifications restrict different objects and disagree sharply on the question of interest, since the median output response to a contractionary shock is positive under Uhlig and negative under ACRR. The rule is also where Uhlig's identification misbehaves. In our common-sample replication the probability that Uhlig's identified set violates the rule signs is 0.72, almost exactly ACRR's published value of 0.76, so most admissible rotations imply a central bank that cuts rates when output or prices rise. The zeros themselves, however, do little within the Uhlig family, because adding only the zeros to Uhlig's restrictions leaves the responses close to Uhlig's, a point ACRR themselves report. The identifying power of their scheme lies in the rule restrictions, and that is what a narrative alternative must recover.

Our Narrative scheme starts from Uhlig's sign restrictions and adds only narrative information and the two qualitative rule signs, with no zero restrictions imposed anywhere. Panel B of Table 1 states the scheme and Panel B of Table 2 lists the episodes. Five signed episodes anchor the shock, and their backing deserves to be spelled out because the defensibility of the identification rests on it. The four contractionary dates are the canonical Romer and

Romer (1989) dates inside the sample. Romer and Romer identified these dates by reading the Federal Reserve's own meeting records for moments when the committee decided to accept output losses in order to reduce inflation, so at each date the direction of the policy impulse is documented in the primary sources rather than inferred from the data. October 1979 is the strongest case, since Volcker's change in operating procedures is the most widely accepted monetary event in the postwar record and is the single episode Antolín-Díaz and Rubio-Ramírez (2018) use in their own monetary application. The restriction imposed at each date is deliberately mild, since it constrains only the sign of one month's structural shock and asserts no magnitude. The expansionary date provides the mirror image from a different information source. On 3 January 2001 the Federal Reserve cut the funds rate by fifty basis points at an unscheduled inter-meeting move, and high-frequency federal funds futures identify it as among the largest easing surprises on record (Kuttner, 2001), so a negative policy shock in that month is documented by market prices rather than by narrative judgment alone.

Two systematic-dominance episodes anchor the rule, and their backing is of a different kind. In December 1988 and March 1984 the economy was booming, the output innovations in the model are large and positive, and the funds rate rose. Reading these tightenings as the rule at work follows the critique of Shapiro (1994) and Leeper (1997) that dates identified from the Federal Reserve's stated intentions capture endogenous responses to inflation rather than exogenous shocks, and the March 1984 episode is especially clean because no narrative source identifies it as a policy shock at all. Growth was running near eight percent and the tightening was universally understood at the time as a response to the boom. Two checks validate this reading. First, 42 percent of ACRR's own benchmark posterior satisfies systematic dominance at both dates, so the restriction sharpens within their credible set rather than contradicting it. Second, the signed episodes are individually mild but jointly powerful. Each passes 20 to 92 percent of Uhlig's admissible rotations on its own, yet jointly they pass about one tenth of one percent, far below the benchmark of roughly one percent that would obtain if violations were independent across dates. Violations are instead systematically correlated across the disinflation dates, meaning that rotations which misread one deliberate disinflation misread them all, which is precisely ACRR's concern about improbable systematic responses expressed through the narrative record.

Table 2: Narrative anchor episodes

Panel A. Oil market. NSR-1 to NSR-5 imposed at each anchor		
date	event	narrative basis
1979:M1	Iranian Revolution	Iranian output collapsed by nearly five million barrels per day while the price more than doubled amid fears of deeper cuts (Kilian, 2009; Hamilton, 2003).
1990:M8	Persian Gulf War	The invasion of Kuwait embargoed seven percent of world output, and the price doubled before collapsing in January 1991 ahead of fundamentals (Kilian, 2009).
2002:M12	Venezuelan strike	The PDVSA strike removed roughly half of Venezuelan output as the Iraq invasion loomed (Kilian and Murphy, 2014).
Panel B. Monetary policy		
date	restriction	narrative basis
1968:M12	shock +	Romer and Romer (1989) disinflation decision.
1974:M4	shock +	Romer and Romer (1989) date, tightening amid the oil crisis.
1978:M8	shock +	Romer and Romer (1989) date.
1979:M10	shock +	Volcker reform, the episode Antolín-Díaz and Rubio-Ramírez (2018) restrict.
2001:M1	shock −	Surprise inter-meeting cut of 3 January 2001 (Kuttner, 2001).
1988:M12	sys. dominance	Tightening in a boom read as the rule at work (Shapiro, 1994; Leeper, 1997).
1984:M3	sys. dominance	Tightening amid eight percent growth, identified by no source as a shock.

Notes. In Panel B, shock + and shock − restrict the sign of the monetary policy shock at the episode date, and systematic dominance requires the funds-rate innovation to be driven mostly by the rule's response to the observed data as defined in Section 2. The Supplemental Appendix documents each episode in full.

The systematic-dominance restrictions are essential rather than decorative, and the reason is visible in equation (2). A signed episode asserts that policy moved more than the rule prescribes, which bounds the systematic response from above at that date. In December 1988 the economy was hot, so a rule with a high output coefficient would classify the observed tightening as systematic and leave no room for a positive shock, and using that episode as a shock date caps ψ_y near 0.3. The easing mirror at January 2001 caps the coefficient from the other side, since a strongly counter-cyclical rule already prescribes deep cuts in a downturn and leaves no room for a negative shock. Shock narratives therefore deliver the zeros and discipline the inflation response, but they cannot by themselves push the output response up, and that is exactly the work the systematic-dominance episodes perform.

Figure 2 reports the posterior of the policy-rule coefficients and contains the heart of the paper. Under Uhlig the rule is diffuse and mostly wrong-signed, with ψ_y centered below zero and with the reserves coefficients spread over roughly $[-0.5, 0.7]$. Under the Narrative scheme the reserves coefficients concentrate at zero without ever being restricted, with ψ_{tr} in $[-0.38, 0.02]$ and ψ_{nbr} in $[-0.00, 0.43]$ at the 16th and 84th percentiles, so ACRR's zero restrictions emerge endogenously from the narrative record. The response coefficients match the benchmark, since we obtain $\psi_y = 0.58$ with interval $[0.24, 1.09]$ against ACRR's 0.85 with interval $[0.26, 2.35]$ and $\psi_p = 2.07$ with interval $[1.17, 3.86]$ against their 2.77 with interval $[0.78, 7.72]$. Both narrative intervals nest inside ACRR's, so the narrative delivers their identification and sharpens it.

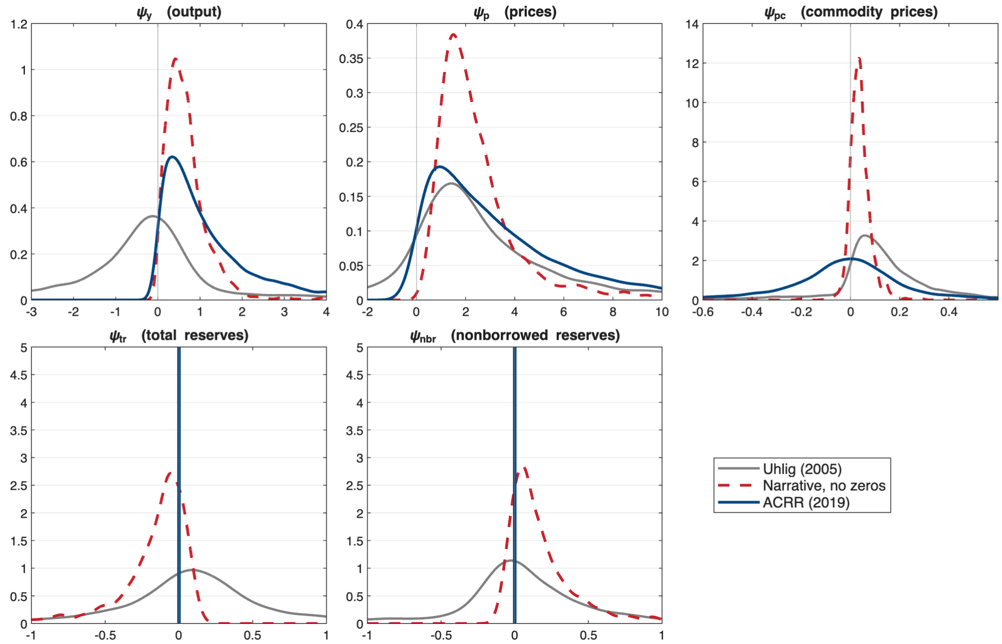

Figure 2: Monetary policy. Posterior densities of the policy-rule coefficients

Notes. The Narrative scheme (red dashed) imposes no zero restrictions, yet the reserves coefficients concentrate at the zero that ACRR (blue vertical line) impose exactly. ACRR also impose $\psi_y, \psi_p \geq 0$, while Uhlig (gray) leaves the rule free.

Figure 3 reports the impulse responses with every posterior draw rescaled so that the impact increase in the funds rate equals 25 basis points. The rescaling matters for the comparison because the narrative selects larger policy shocks than the benchmark, so raw response gaps would mix the size of the shock with its transmission. Per 25 basis points of tightening, the Narrative scheme and ACRR are nearly coincident on output, prices, reserves, and the funds rate, with median gaps of 10 to 25 percent of the response amplitude against 25 to 250 percent for Uhlig, and output falls with high posterior probability, so Uhlig's output puzzle is resolved by the narrative record rather than by zeros. The one remaining gap is the commodity-price response, which is inherited from the Uhlig base restrictions rather than from the narrative. As in the oil application the identification is tight, since characterizing the surviving set required forty-five million posterior draws with the narrative screened inside the sampler, of which about twenty-nine hundred effective draws survive. The Supplemental Appendix reports the sensitivity of these results to the treatment of December 1988, showing that the variant which keeps it as a shock date still recovers the zeros and the inflation response but caps the output coefficient, together with all intermediate specifications.

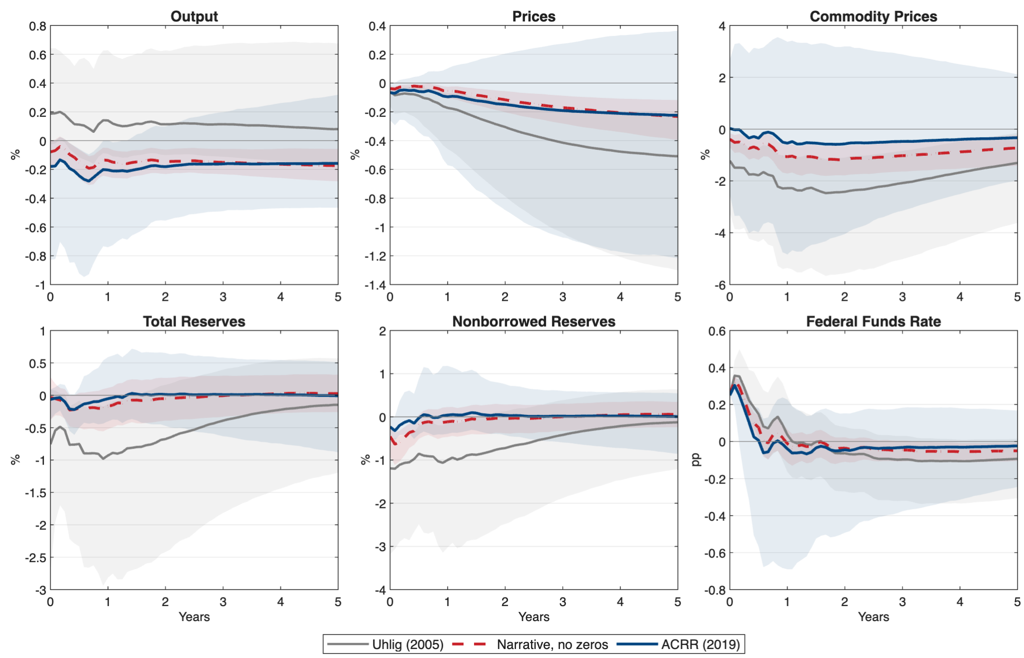

Figure 3: Monetary policy. Impulse responses to a contractionary policy shock

Notes. Every posterior draw is rescaled so that the impact increase in the federal funds rate equals 25 basis points. Shaded regions are 68 percent pointwise credible bands.

5 Conclusion

A central concern in the set-identified SVAR literature is that sign restrictions alone admit a wide set of rotations, yielding diffuse posteriors and fragile conclusions. The common response has been to impose harder restrictions, a tight elasticity bound in the oil market and exact zero restrictions on the policy rule in monetary VARs, in order to stabilize the system. We take a different route in both settings and tie identification to well-documented historical episodes, letting the narrative record discipline the rotation set.

The two applications deliver the same verdict. In the oil market, removing the elasticity bound leaves impulse responses, variance decompositions, and historical decompositions essentially unchanged once the narrative is imposed, and the narrative eliminates the implausibly elastic models that sign restrictions alone admit. In the monetary VAR, removing the zero restrictions leaves the benchmark intact once the narrative is imposed, the reserves coefficients concentrate at zero endogenously, the rule coefficients match the benchmark, and Uhlig's output puzzle is resolved. In both cases very few rotations survive the narrative, indicating that the discipline comes from the historical record rather than from the hard restrictions.

The monetary application also clarifies what narrative information can and cannot do, which we regard as practical guidance for applying the method elsewhere. Restrictions on shocks assert that policy or supply moved more than the systematic forces prescribe, so they bound systematic responses from above and can never establish that a response is positive. Recovering a full behavioral rule therefore requires narrative information about the rule itself, which our systematic-dominance restriction provides by pointing to episodes where the observed movement was the system working as intended. A researcher applying the method should thus select episodes of both kinds, document each with sources whose interpretation is not in dispute, impose the restrictions jointly, and read the tightness of the surviving rotation set as a diagnostic of how informative the history is. Carefully chosen, historically grounded narrative restrictions offer a transparent and credible compromise in the stabilization debate, cutting down the admissible rotation set and delivering stable conclusions while avoiding the contentiousness of tight elasticity bounds and exact zero restrictions.

References

Antolín-Díaz, Juan and Juan F. Rubio-Ramírez. 2018. "Narrative Sign Restrictions for SVARs." *American Economic Review* 108 (10):2802–2829.

Arias, Jonas E, Dario Caldara, and Juan F Rubio-Ramírez. 2019. "The systematic component of monetary policy in SVARs. An agnostic identification procedure." *Journal of Monetary Economics* 101:1–13.

Baumeister, Christiane and James D. Hamilton. 2019. "Structural Interpretation of Vector Autoregressions with Incomplete Identification: Revisiting the Role of Oil Supply and Demand Shocks." *American Economic Review* 109 (5):1873–1910.

Baumeister, Christiane and Gert Peersman. 2013. "The Role of Time-Varying Price Elasticities in Accounting for Volatility Changes in the Crude Oil Market." *Journal of Applied Econometrics* 28 (7):1087–1109.

Caldara, Dario, Michele Cavallo, and Matteo Iacoviello. 2019. "Oil Price Elasticities and Oil Price Fluctuations." *Journal of Monetary Economics* 103:1–20.

Hamilton, James D. 2003. "What Is an Oil Shock?" *Journal of Econometrics* 113 (2):363–398.

Herrera, Ana María and Sandeep Kumar Rangaraju. 2020. "The Effect of Oil Supply Shocks on US Economic Activity: What Have We Learned?" *Journal of Applied Econometrics* 35 (2):141–159.

Kilian, Lutz. 2009. "Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market." *American Economic Review* 99 (3):1053–1069.

Kilian, Lutz and Thomas K. Lee. 2014. "Quantifying the Speculative Component in the Real Price of Oil: The Role of Global Oil Inventories." *Journal of International Money and Finance* 42:71–87.

Kilian, Lutz and Daniel P. Murphy. 2012. "Why Agnostic Sign Restrictions Are Not Enough: Understanding the Dynamics of Oil Market VAR Models." *Journal of the European Economic Association* 10 (5):1166–1188.

———. 2014. "The Role of Inventories and Speculative Trading in the Global Market for Crude Oil." *Journal of Applied Econometrics* 29 (3):454–478.

Kilian, Lutz and Xiaoqing Zhou. 2018. "Modeling Fluctuations in the Global Demand for Commodities." *Journal of International Money and Finance* 88:54–78.

Kuttner, Kenneth N. 2001. "Monetary policy surprises and interest rates. Evidence from the Fed funds futures market." *Journal of Monetary Economics* 47 (3):523–544.

Leeper, Eric M. 1997. "Narrative and VAR approaches to monetary policy. Common identification problems." *Journal of Monetary Economics* 40 (3):641–657.

Romer, Christina D and David H Romer. 1989. "Does monetary policy matter? A new test in the spirit of Friedman and Schwartz." In *NBER Macroeconomics Annual 1989, Volume 4*. MIT Press, 121–170.

Shapiro, Matthew D. 1994. "Federal Reserve policy. Cause and effect." In *Monetary Policy*, edited by N Gregory Mankiw. University of Chicago Press, 307–334.

Uhlig, Harald. 2005. "What are the effects of monetary policy on output? Results from an agnostic identification procedure." *Journal of Monetary Economics* 52 (2):381–419.